引用格式：洪永淼, 汪寿阳. 人工智能新近发展及其对经济学研究范式的影响. 中国科学院院刊, 2023, 38(3): 353-357, doi: 10.16418/j.issn.1000-3045.20230303001.
Hong Y M, Wang S Y. Impacts of cutting-edge artificial intelligence on economic research paradigm. Bulletin of Chinese Academy of Sciences, 2023, 38(3): 353-357, doi: 10.16418/j.issn.1000-3045.20230303001. (in Chinese)

人工智能新近发展及其对经济学研究范式的影响

洪永淼[1,3,4]　汪寿阳[1,2,3*]

1 中国科学院数学与系统科学研究院　北京　100190
2 中国科学院预测科学研究中心　北京　100190
3 中国科学院大学　经济与管理学院　北京　100190
4 中国科学院大学　数字经济监测预测预警与政策仿真教育部哲学社会科学实验室（培育）　北京　100190

摘要　以ChatGPT及其大语言模型为代表的人工智能将对经济学研究范式产生深远影响。目前，经济学实证研究一般使用参数维数较少、经济可解释性较强的小模型。然而，经济金融系统是一个超高维动态复杂系统，受多种因素的影响，且这些因素之间的关系呈现非线性与时变性特征，小模型无法刻画其本质规律。大模型可有效减少系统偏差，更好刻画复杂经济系统的特征与运行规律；而海量数据的使用可避免模型过度拟合，使大模型具有较好的泛化能力即样本外预测能力。为支持经济学及社会科学其他领域的大模型的估计、推断与预测，需要利用人工智能技术整合各种异构、异源、异频数据，构建大规模计量经济学数据库，并加强大算力等信息技术基础设施建设。目前，ChatGPT及大模型等前沿人工智能技术仍存在局限性，如：无法像人类一样进行批判性思考或想象，只有预测能力；基于大数据的人工智能因果推断本质上是一种统计关系推断，需要引入经济理论或实验方法帮助识别真正的因果关系；人工智能技术不能改变经济学实证研究从样本推断总体性质的本质；同时，由于互联网大数据存在大量虚假信息或样本选择偏差等问题，基于人工智能所获得的结论的可靠性需要验证。

关键词　人工智能，ChatGPT，大模型，维数灾难，因果推断

DOI　10.16418/j.issn.1000-3045.20230303001

在数智时代，大数据与人工智能改变了经济主体行为与经济运行方式，也改变了经济学研究范式。大数据为人类提供了认识与改造世界的新思维，即大数据思维，这是通过大数据发现复杂系统的运行规律、

* 通信作者
资助项目：国家自然科学基金基础科学中心项目（71988101）
修改稿收到日期：2023年3月7日

解决现实问题、预测未来变化的新范式；大数据思维的实现方式主要基于人工智能技术与方法。现代经济学的主流研究范式是实证研究，即以数据为基础推断经济变量之间的逻辑关系，特别是因果关系，从而揭示经济运行规律；计量经济学是其最主要的方法论。大数据的可获得性进一步强化了这种研究范式。大数据与人工智能催生了数据驱动研究范式，计算社会科学也因此应运而生。那么，人工智能前沿技术特别是新近诞生的聊天机器人 ChatGPT 及其基础模型技术的发展将如何改变经济学乃至社会科学的研究范式?

1 ChatGPT 及其大语言模型方法论

近几年来，人工智能特别是利用人工智能技术生成内容的技术获得迅猛发展，其中一个代表性技术进步是通用型聊天机器人 ChatGPT 的成功开发并投入使用。ChatGPT 是一种基于互联网文本数据进行训练的文本生成深度学习模型，并通过强化学习及人机对话的方式持续提供反馈，能够较好执行各种自然语言处理任务。ChatGPT 是人工智能特别是自然语言处理领域的一个革命性技术突破，代表人工智能发展的一个新方向，预计对人类生产方式、生活方式及社会治理方式等各个领域将产生深远影响。

ChatGPT 的主要方法论是“规模至上”，其算法基础是大语言模型。大模型是参数维数极大的模型，这些参数需要通过数据训练或估计。大语言模型是指输入数据主要为文本数据的大模型。大语言模型在深度学习发展阶段就已出现。随着人工神经网络模型的隐藏层不断增加，其参数数量呈现快速增长。文本数据本质上是高维或超高维数据，简约模型无法刻画异质性高维数据的特征，因此需要使用大语言模型。从计量经济学与统计学视角看，大模型的最大优点是系统偏差比较小。同时，由于 ChatGPT 的训练数据主要来自互联网文本数据，样本容量极其庞大，从而保证了大语言模型参数的估计精度。因此，大语言模型具有比较强的泛化能力，即样本外预测能力。特别是大语言模型具有非线性规模效应，随着规模的增加，其预测能力呈现非线性增加。

在经济管理领域，人类很多决策均基于样本外预测。如果使用某个领域的文本数据来训练模型，模型维度可能不用太大。但是，对通用人工智能技术，随着异质性文本数据的大量增加，小模型偏差较大，其样本外预测能力变差。这时需要扩大模型规模，使用更多数据训练模型，以提高泛化能力。随着计算机科学特别是算力与算法的快速进步，这种不断扩大规模的方法已成为通用人工智能技术的一个发展模式。

长期以来，计量经济学与统计学面临的一个困扰是“维数灾难”：当模型参数维数相比数据容量不是很小时，虽然模型偏差比较小，但因为参数维数大，在有限数据容量条件下，对每个参数的估计不甚精准，导致模型过度拟合，其样本外预测能力较差。“维数灾难”不仅存在于计量经济学与统计学，在社会科学和自然科学很多领域也存在。ChatGPT 通过使用海量互联网大数据，确保训练数据容量远大于模型参数维数，从而避免了“维数灾难”。另外，大语言模型的结构设计是深度、多维、多层、多头的注意力结构，可自适应高维空间的稀疏与不平衡数据结构，为避免“维数灾难”提供了一种表示学习的解决方案。

作为学术研究的好助手，ChatGPT 可大幅提高研究效率。在经济学研究中，可借助 ChatGPT 搜索信息、收集数据、撰写文献综述、编写代码、检查程序、设计实验方案、翻译文本等。虽然目前 ChatGPT 的表现有不少缺点，其整体智能水平与人类相比尚有不小的差距；但可以预计，随着人工智能技术的发展以及 ChatGPT 广泛应用的经验积累，ChatGPT 的智能程度在很多方面将日益接近甚至超过人类。

ChatGPT 及其大语言模型方法论可能会对经济学

乃至整个社会科学的研究范式产生深远影响。例如，实验经济学研究表明，人类经济行为并不满足完全理性假设，更多表现为有限理性，甚至存在预期偏差。ChatGPT 及其大语言模型能够改进理性经济人假设；因此，基于人工智能或由人工智能辅助的经济决策可提供更有效的资源配置方案、改进政策评估精准性，以及提升经济决策的科学性等。本文聚焦讨论 ChatGPT 及其大语言模型方法论对经济学研究范式的可能影响。

2 大模型与经济学研究范式

2.1 计量经济学小模型范式

过去 40 年，经济学研究范式发生了一个深刻变革，即所谓的“实证革命”，也被称为“可信性革命”。作为经济学实证研究最主要的方法论，计量经济学也日益成为社会科学很多领域的主流研究范式。2021 年诺贝尔经济学奖得主 Joshua Angrist 曾指出，应用计量经济学所考虑的问题和其他社会科学或者流行病学所考虑的问题并无本质区别，任何希望从数据中得到有用推断的人都可称为应用计量经济学家。

计量经济学与统计学的一个基本建模原则是，使用尽量简约的模型即小模型刻画数据特征及变量之间的关系。在 20 世纪，统计学存在参数与非参数建模的方法论之争。统计学家 Ronald Fisher 认为非参数模型因其参数维数高而估计不精确，主张使用参数维数较少的参数模型；另一位统计学家 Karl Pearson 则关注参数模型可能误设而产生较大偏差，主张使用非参数模型。计量经济学也有类似争论。2003 年诺贝尔经济学奖得主 Robert Engle 主张从特定模型出发，通过检验遗漏变量等计量经济学方法拓展模型，这是所谓的“从特殊到一般”的建模方法；而伦敦计量经济学派代表人物 David Hendry 则主张从高维的一般模型出发，通过统计学假设检验与经济理论约束条件等方法得到特定模型，这是所谓的“从一般到特殊”的建模方法。“从一般到特殊”的建模方法更适合大数据分析，更接近数据驱动研究范式。但是，两种建模方法的最终目的都是获得一个具有经济可解释性的简约模型，参数不多且有经济含义，同时拥有良好的样本外预测能力。

为了获得简约模型，计量经济学与统计学提出了很多经典降维方法，如主成分分析、因子模型、模型选择、经济理论约束等。在大数据时代，由于潜在解释变量或预测变量很多，新的统计降维方法也不断产生。例如，人工智能与统计学的交叉产生了一个新领域——统计学习，其中一个代表性降维方法是 LASSO 回归。LASSO 的基本思想是假设大量潜在解释变量中只有少数变量有重要影响，在此稀疏性假设下，通过引入适合的惩罚项，以牺牲估计偏差为代价，换取估计方差的大幅度减少，从而显著降低均方误差，达到精准选择重要变量和改进样本外预测的目的。

2.2 计量经济学应发展大模型范式

长期以来，计量经济学与统计学使用的数据均是“小”数据，大量数据信息没有被利用。计量经济学与统计学的研究范式以模型为重心，以降维为导向，这主要归因于经济学研究的可解释性要求、数据容量有限及计算资源约束。随着数智时代的到来，人工智能特别是机器学习，如 LASSO、决策树、随机森林、支持向量机、人工神经网络模型及深度学习等方法，被广泛应用于经济学及社会科学其他领域的研究。实证研究表明，在经济金融预测方面，机器学习显著优于传统计量经济学模型，但尚未达到令人满意的程度，尤其是与其他领域（如人脸识别）相比，预测精准性仍有待提升。目前机器学习在这方面表现较差的主要原因是经济金融系统是一个超高维动态复杂系统，影响因素繁多并互相关联，其相互关系呈现非线性特征并有时变性；此外，经济金融系统还受经济主体心理因素（如预期、情绪等）的影响。如此复杂的

经济系统是不可能用小模型来刻画其本质规律的。

为了显著改进经济金融预测与提升模型的解释力，可在计量经济学建模中考虑大模型范式。这里所说的大模型是指参数维度较大的计量经济学模型，不一定要使用互联网海量文本数据或深度学习方法，模型参数维度也可远小于 ChatGPT 那样的数量级。大模型可容纳高维潜在的影响因素，允许变量之间存在非线性关系，允许模型参数具有时变性，因此能够显著减少模型偏差。计量经济学与统计学的模型组合方法就是一种大模型范式，可得到比较稳健、精准的样本外预测。经济学与计量经济学已有一些大模型，如投入产出模型、可计算一般均衡模型、宏观经济学联立方程模型，以及 1980 年诺贝尔经济学奖得主 Lawrence Klein 提出的世界连接模型等，但这些模型大多受到数据容量等的限制。

大模型方法能够成功的关键是适用于计量经济学建模的大规模数据的可获得性。目前，计量经济学模型所使用的数据本质上大多是“小”数据，不足以支持大模型的估计、推断与预测。如何有效利用现有各种结构化和非结构化数据，如何利用人工智能技术整合各种类型、各种抽样频率、各种来源的数据，构建大规模的计量经济学数据库，是大模型方法的一个前提条件。此外，大算力等信息技术基础设施建设对大模型在计量经济学中的应用也不可或缺。在国家层面，即将成立的国家数据局将负责协调推进数据基础制度建设，统筹数据资源整合共享和开发利用，这将为大模型范式的探索与推广提供重要支持。

3 人工智能技术的局限性

人工智能特别是 ChatGPT 及其大语言模型正在推动经济学研究范式的深刻变革，但是人工智能技术也存在局限性。

（1）以 ChatGPT 为代表的前沿人工智能目前仍没有人类的意识或理解能力，只有预测能力。ChatGPT 不久前在接受美国《时代》周刊采访时表示，“说我有知觉或意识是不准确的。作为一个大型语言模型，我只是一个机器学习模型，没有和人类一样的意识。我没有思想、感觉或经历，没有能力做决定或独立判断。” ChatGPT 无法理解生成单词的上下文语境或含义，只能根据给定的训练数据，根据某些单词或单词序列一起出现的概率生成文本。因此，其不大可能拥有与人类一样的批判性思维和想象力，无法从现有数据推断出重大创新性成果。

（2）基于大数据的人工智能因果推断本质上是一种统计关系推断，并不一定是真正的因果关系。所谓因果关系，是指其他因素不变的条件下，某一变量的变化必然引起另一变量的变化。识别因果关系是经济学乃至社会科学研究的最根本问题。经济学因果推断的最有效方法是实验方法，如随机控制实验。在因果推断方面，计量经济学与统计学取得了长足的方法论进展。2021 年，经济学家 Joshua Angrist、David Card 和 Guido Imbens 由于他们在因果推断方法论方面的贡献而获得诺贝尔经济学奖。但是，由于现实大数据基本上是观测数据而非实验数据，基于人工智能的因果推断本质上是一种相关关系或预测关系，是两个变量之间在计算机中的统计关联。在某种意义上，基于大数据的人工智能因果关系类似于英国哲学家休谟所阐述的因果关系，即两个现象或变量在人脑中形成的惯常联系。因此，人工智能因果关系和经济学因果关系并不完全一致。要识别经济学因果关系，不能仅依靠人工智能因果推断技术，还必须有经济理论的指导或者引入实验方法。这一点，对可解释性较弱的大模型尤为重要。

（3）人工智能没有改变经济学实证研究的本质，即从样本推断总体性质。虽然以 ChatGPT 为代表的人工智能使用了几乎整个互联网文本数据，但互联网大数据并非全样本。人类经济社会发展是一个长期历史过程，现有互联网大数据只是这个历史过程的一个样

本，即便其容量特别庞大。此外，经济发展在不同历史阶段具有不同的特征，经济运行规律因此会呈现出时变性——有时是缓慢变化，有时是突变，这使得人工智能对经济金融未来变化趋势的预测更具挑战性。

（4）模型、算法与数据的可靠性有待验证。人工智能正在推动经济学乃至整个社会科学研究从模型驱动范式转变到数据驱动范式，从数据直接获得稳健的结论，克服了模型驱动范式得到的结论可能会因模型改变而变化的缺陷。但是，互联网大数据存在各种虚假信息，以 ChatGPT 为代表的人工智能仍无法确认其表述内容的真实性。此外，虽然容量非常庞大，互联网大数据也存在“样本选择偏差”问题，“数字鸿沟”就是一个重要例子。这些问题不可避免地会影响基于互联网大数据的 ChatGPT 乃至人工智能所获得的结论的可靠性与科学性。

4 展望

我们正处于大数据、大科技、大模型的时代，加上中国超大人口规模和超大经济规模给人工智能技术带来的广阔应用场景，这些将为经济学研究提供大量丰富的素材。应当充分利用所有数据资源，积极探索大模型研究范式，以更好揭示中国复杂经济系统的运行与发展规律。

需要指出，强调大模型并不意味小模型不重要，大小模型分别适用于不同情境；强调大数据并不意味“小”数据不重要，“小”数据的信息密度通常更高；强调文本数据等非结构化大数据并不意味结构化数据不重要。历史上，热力学与统计物理学的发展经验或许可以提供一些有用启示：物理学黑体辐射研究曾经出现两个理论——在短波范围拟合较好的“维恩近似”，以及在长波范围拟合较好的“瑞利-金斯定律”；后来，普朗克将两者综合起来，提出了覆盖全波段范围的“黑体辐射定律”。

在经济学研究中，大、小模型各有优缺点，是否可以找到更好的科学方法把这两种模型结合在一起，从而提高模型的可解释性和预测力？关于小模型的研究已有很长的历史，经济学家对小模型在经济学实证研究中的优点与局限性已有比较深刻的了解。但是，关于大模型的研究及大模型在经济学实证研究中的应用还很少，因此需要大力探索发展大模型方法，并与小模型方法进行比较。

洪永淼　中国科学院数学与系统科学研究院关肇直首席研究员，中国科学院大学经济与管理学院、中国科学院大学数字经济监测预测预警与政策仿真教育部哲学社会科学实验室（培育）特聘教授。发展中国家科学院院士，世界计量经济学会会士。研究领域为计量经济学、时间序列分析、金融计量学、统计学、中国经济等。E-mail: ymhong@amss.ac.cn

汪寿阳　中国科学院数学与系统科学研究院、中国科学院预测科学研究中心特聘研究员，中国科学院大学经济与管理学院特聘教授。发展中国家科学院院士，国际系统与控制科学院院士。研究领域为决策分析、供应链管理、风险管理、经济分析与预测、计量经济与政策分析等。E-mail: sywang@amss.ac.cn

■责任编辑：岳凌生